Statistical analysis of neural data (GR8201)

Fall 2022

This is a Ph.D.-level topics course in statistical analysis of neural data. Students from statistics, neuroscience, and engineering are all welcome to attend. A link to the previous iteration of this course is here.

Time: W 1:30-3
Place: JLG L5-084
Professor: Liam Paninski; Office: Zoom. Email: **liam at stat dot columbia dot edu**. Hours by appointment.

Prerequisite: A good working knowledge of basic statistical concepts (likelihood, Bayes' rule, Poisson processes, Markov chains, Gaussian random vectors), including especially linear-algebraic concepts related to regression and principal components analysis, is necessary. No previous experience with neural data is required.
Evaluation: Final grades will be based on class participation and a student project. Additional informal exercises will be suggested, but not required. The project can involve either the implementation and justification of a novel analysis technique, or a standard analysis applied to a novel data set. Students can work in pairs or alone (if you work in pairs, of course, the project has to be twice as impressive). See this page for some links to available datasets; or talk to other students in the class, many of whom have collected their own datasets.
Course goals: We will introduce a number of advanced statistical techniques relevant in neuroscience. Each technique will be illustrated via application to problems in neuroscience. The focus will be on the analysis of single and multiple spike train and calcium imaging data, with a few applications to analyzing intracellular voltage and dendritic imaging data. Note that this class will not focus on MRI or EEG data. A brief list of statistical concepts and corresponding neuroscience applications is below.

Statistical concept / technique	Neuroscience application
Point processes; conditional intensity functions	Neural spike trains; photon-limited image data
Time-rescaling theorem for point processes	Fast simulation of network models; goodness-of-fit tests for spiking models
Bias, consistency, principal components	Spike-triggered averaging; spike-triggered covariance
Generalized linear models	Neural encoding models including spike-history effects; inferring network connectivity
Regularization; shrinkage estimation	Maximum a posteriori estimation of high-dimensional neural encoding models
Laplace approximation; Fisher information	Model-based decoding and information estimation; adaptive design of optimal stimuli

Mixture models; EM algorithm; Dirichlet processes	Spike-sorting / clustering
Optimization and convexity techniques	Spike-train decoding; ML estimation of encoding models
Markov chain Monte Carlo: Metropolis-Hastings and hit-and-run algorithms	Firing rate estimation and spike-train decoding
State-space models; sequential Monte Carlo / particle filtering	Decoding spike trains; optimal voltage smoothing
Fast high-dimensional Kalman filtering	Optimal smoothing of voltage and calcium signals on large dendritic trees
Markov processes; first-passage times; Fokker-Planck equation	Integrate-and-fire-based neural models
Hierarchical Bayesian models	Estimating multiple neural encoding models
Amortized inference	Spike sorting; stimulus decoding

For those new to neuroscience: While we will cover all the necessary background as we go, for those who want to explore the material in greater depth, there are a bunch of good computational neuroscience resources. The recent Neuromatch Academy is a good place to start. A very non-exhaustive list of useful books (each of which emphasize different topics, albeit with some overlap): Theoretical Neuroscience, by Dayan and Abbott; Spiking Neuron Models, by Gerstner et al; and Spikes: exploring the neural code, by Rieke et al. The first chapter of the Spikes book has been kindly made available online - this makes a nice overview of some of the questions we will address in this course. The full text of the Gerstner et al book is online. Another good online tutorial is available here.

A couple good older online courses in computational neuroscience: one directed by Raj Rao and Adrienne Fairhall, and another by Wulfram Gerstner.

For those new to statistics: The new book by Kass et al is an excellent introduction to statistics, illustrated with a number of neural examples; Columbia e-link here. Also, here is an excellent online book on convex optimization. Finally, Cox and Gabbiani have written a nice Matlab-based book on Mathematics for Neuroscientists, available online here if your library has access. A lot of very useful background material, along with some more advanced ideas.

Schedule

Date	Topic	Reading	Notes
Sept 7	Intro and overview	Paninski and Cunningham, `18; International Brain Lab, '17	

Sept 14	Signal acquisition: spike sorting	Lewicki '98; Pachitariu et al '16; Lee et al '20; Steinmetz et al '21; Calabrese and Paninski '11, Boussard et al '21, Varol et al '21, Wang et al '19, Zanos et al '11	EM notes; Blei et al review on variational inference. Guest lecture by Julien Boussard and Charlie Windolf.

Thanks to the NSF for support.